Methodology I: ANALYSIS

Notes for CS 6400: Database Systems Concepts & Design
Georgia Tech (Dr. Jay Summet with Dr. Leo Mark Christensen), Summer 2017
as recorded by Brent Wagenseller

Database Application Development Methodology - Assumptions

- The **Database Application Development Methodology** was specifically designed for applications supported by databases – it will NOT work for general SDP projects
- The Database Application Development Methodology assumes that
  - Business processes are well-designed
  - Documents are known
    - Documents are anything that is input to – or output from – anything that runs on the DB
  - Tasks are known
    - The processing that takes place, using
  - System boundaries are known
  - One database schema unifying all views can be designed
    - This is a VERY difficult task, due to interests, goals, power, politics
    - There could be problems with the methodology or organization
      - Answer: its ALWAYS the organization

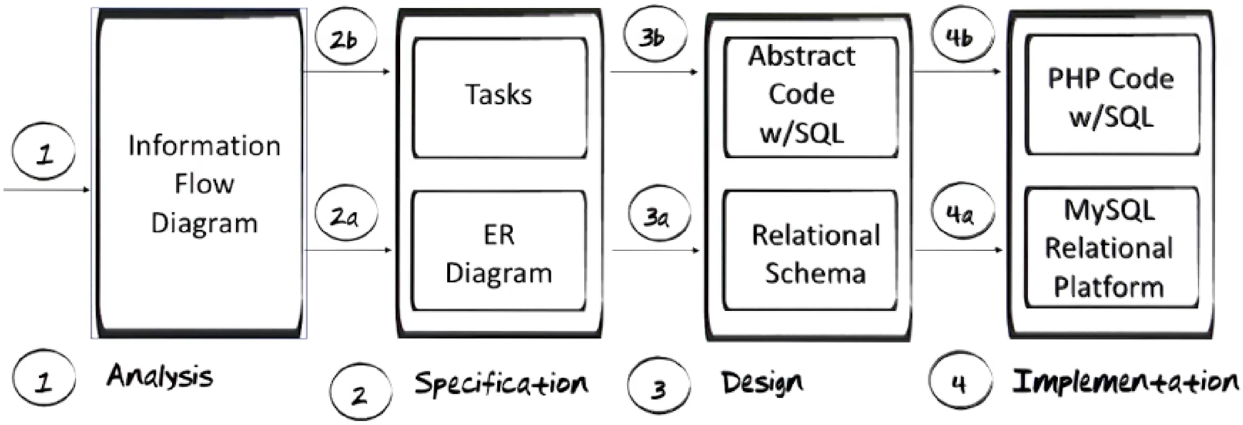
The Software Process

- Recall the SDP Waterfall process has these steps:
  - Business Process (re-)design
  - Analysis
  - Specification
  - Design
  - Implementation
  - Testing
  - Operation
  - Maintenance
- <u>That said, for the Database Application Development Methodology we will only look at:</u>
  - **Analysis**
  - **Specification**
  - **Design**
  - **Implementation**

Overview of the Methodology: Data First!

- <u>In the SDP, the *process* comes first; however, in the Database Application Development Methodology, *data* comes first!</u>

    - That is to say, the process – and everything else – is centered around the data
- The Four Phases of the Methodology:



    - Analysis Phase
        - In the analysis phase, we start with all of the information provided by the customer (customer requirements), and the end product of the phase is an **Information Flow Diagram**
    - Specification
        - The Information Flow Diagram is the input to the specification process
        - In the specification process, the first thing we do is create an ER Diagram
        - With that, we then create a specification of the **tasks** / applications that need to run on the database represented by the **ER / EER diagram**
    - Design
        - We first concentrate on translating the ER Diagram to a **relational database schema**
        - With this in place, we take the tasks that represent what needs to happen in the ER diagram (from Specification) and represent them in **abstract code** accessing the relational database represented by the schema
    - Implementation
        - We take the relational database schema and we use a specific implementation to create the **relational database** (in our examples we will be using MySQL)
        - We then look at the abstract code with the embedded SQL (PHP) to represent the tasks needed to run on this database
- If we do not know the customer's requirements, the methodology will not result in a database application that meets the requirements the user might have

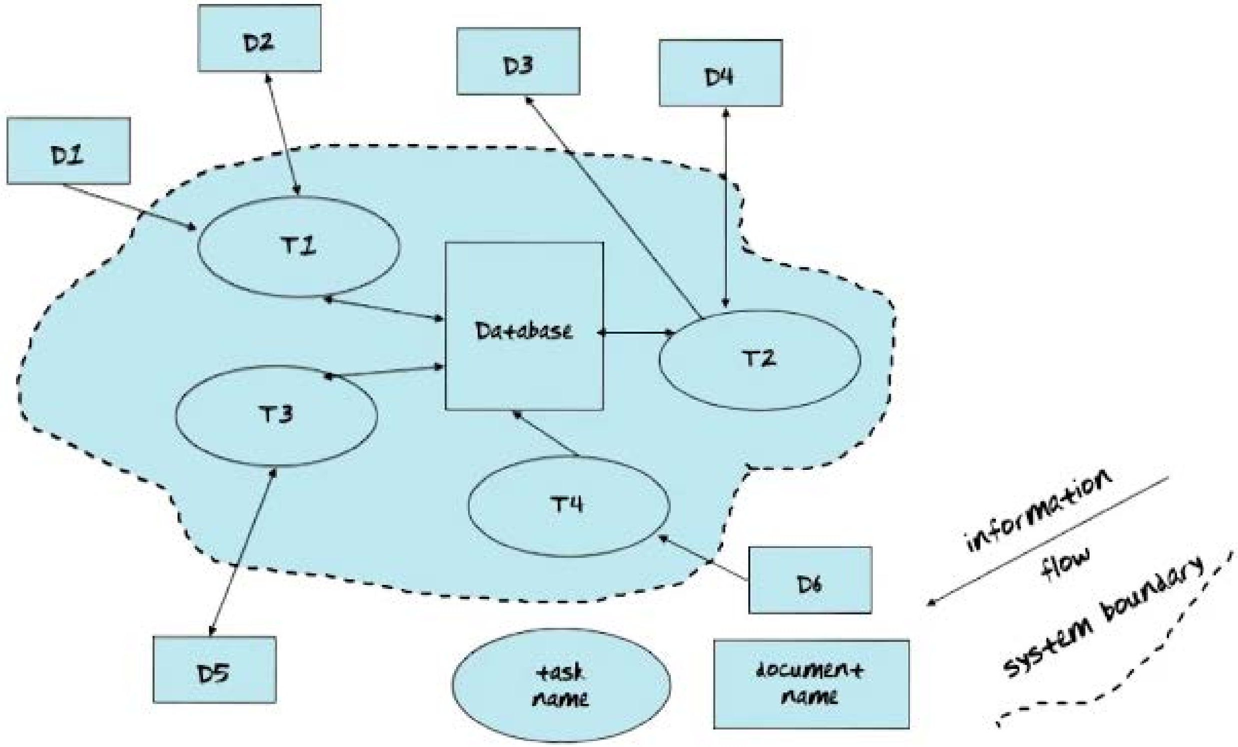
Example Project Description

- An example will be laid out that we can follow
- We will assume the assumptions – business process, documents, tasks, and system boundaries – are all laid out for us

- The example will be “GTOnline”, a simple networking application similar to facebook and linkedin
- BRENTS NOTE: There is a case study that we must download entitled “GTPEgtonline_description.pdf”

Analysis

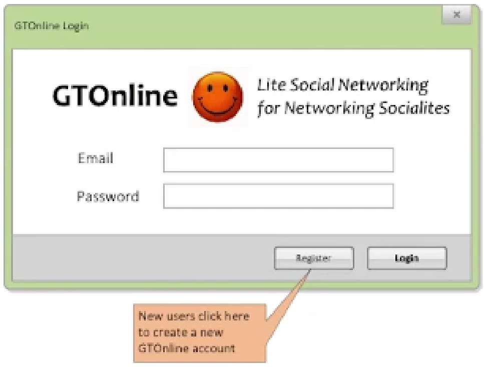
- In the analysis, we take the customers requirement and create an information flow diagram
- **Information Flow Diagram**



  - The ellipses represent tasks
  - The rectangles represent input / output screens
  - One task can be associated with multiple documents
  - The arrows represent information flow
    - Note this is NOT a flow chart; rather, it’s a chart that shows potential information flow
    - BRENTS NOTE: That said the arrows look specific, so you may want to be deliberate with them
  - The broken line is the system boundary
  - General rules
    - Again – its an information flow, NOT a control flow
    - NEVER connect two documents
    - NEVER connect two tasks

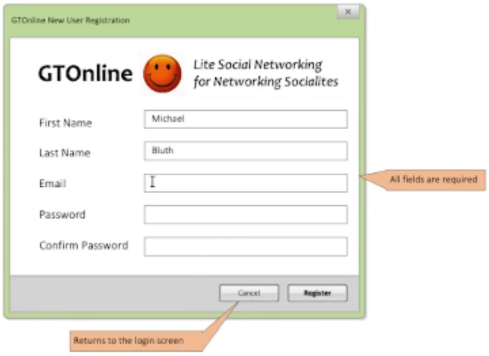
Examples of Requirements

- Note that the picture is referred to the 'document'
  - BRENTS NOTE: Is this what is meant by 'document'? Will we have to have mock screens in place for the first part of the project?


GTOnline Login
×
GTOnline
Lite Social Networking
for Networking Socialites
Email
Password
Register
Login
New users click here to create a new GTOnline account

Requirement: ...Logging In to GTOnline via the login screen. All users are uniquely identified by his or her Email Address. Providing a valid Email Address and Password combination will log the user into the system. Providing invalid login credentials should display an error message and return the user to the login screen.

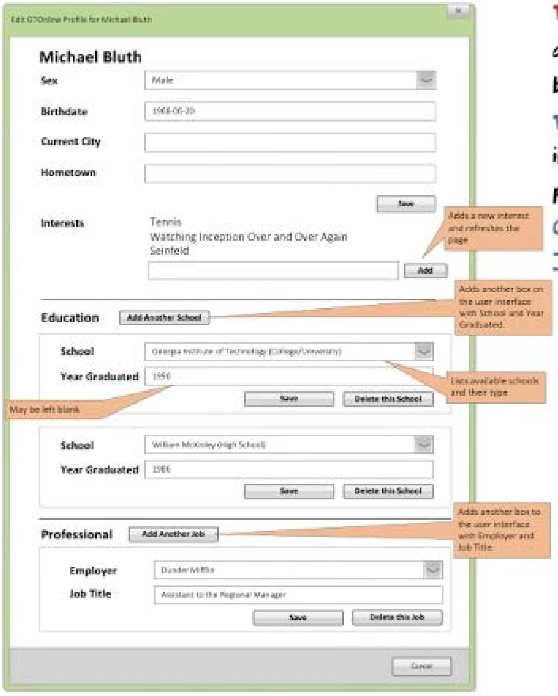
  - This is purely an **input document**, as it ONLY take an input



First Name: Michael

Last Name: Bluth

Requirement: ...Users who are new to GTOnline must register first. A Register button is provided directly on the login page. Clicking this button displays the new user registration form. After the user clicks Register, the system should verify that all fields are filled in, that the Email Address has not already been registered, and that the Password and Confirm Password fields are equal.


Edit GTOnline Profile for Michael Bluth
×
Michael Bluth
Current City
Hometown
Save
Add
Adds a new interest and refreshes the page
Education
Add Another School
Adds another box on the user interface with School and Year Graduated
Save
Delete this School
Lists available schools and their type
May be left blank
Save
Delete this School
Professional
Add Another Job
Adds another box to the user interface with Employer and Job Title
Save
Delete this Job
Cancel

Sex: Male

Birthdate: 1968-06-20

Interests: Tennis
Watching Inception Over and Over Again
Seinfeld

School: Georgia Institute of Technology (College/University)

Year Graduated: 1990

School: William McKinley (High School)

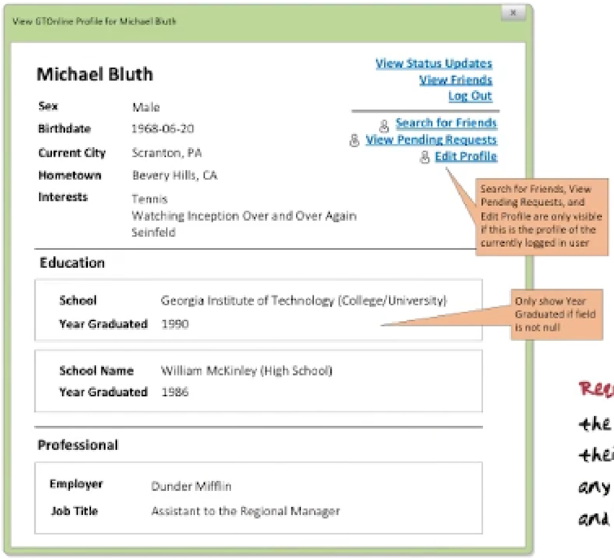
Year Graduated: 1986

Employer: Dunder Mifflin

Job Title: Assistant to the Regional Manager

Requirement: ...All GTOnline users (except administrative users) have a User Profiles containing basic information about them. After a new user Registers with GTOnline, they should be taken immediately to the Edit Profile screen. The basic profile properties include the user's Sex, Birthdate, Current City, Hometown, and any number of Interests.

- Note that this document also pulls data, as its needed to populate the dropdowns


View GTOnline Profile for Michael Buth

Michael Bluth

Sex	Male
Birthdate	1968-06-20
Current City	Scranton, PA
Hometown	Bevery Hills, CA
Interests	Tennis Watching Inception Over and Over Again Seinfeld

Education

School	Georgia Institute of Technology (College/University)
Year Graduated	1990

School Name	William McKinley (High School)
Year Graduated	1986

Professional

Employer	Dunder Mifflin
Job Title	Assistant to the Regional Manager

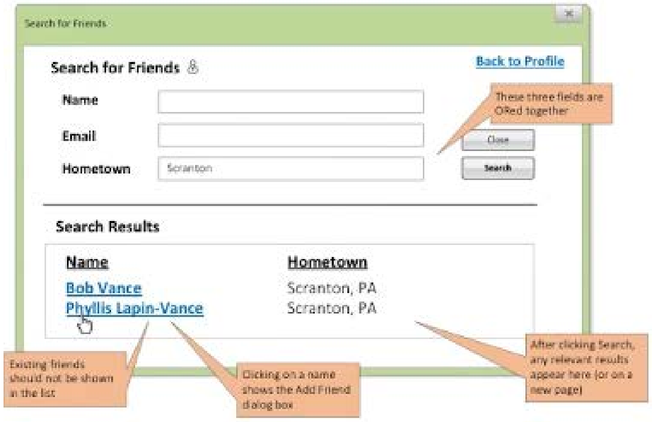
Search for Friends, View Pending Requests, and Edit Profile are only visible if this is the profile of the currently logged in user

Only show Year Graduated if field is not null

Requirement: ...The profile also contains Professional information. The user will select his or her Employer from the list and then provide a Job Title

Requirement: ...The profile also contains information about the user's Education. The set of available Schools and their Types is maintained by GTOnline. A user can have any number of schools associated with his or her profile and can also provide a Graduation Date for each school.

- This ‘view user’ document is an **output document** – it does NOT write anything to the database
  - NOTE: The control information (links in the top right) do not count towards this


Search for Friends

Search for Friends

Name	
Email	
Hometown	Scranton

These three fields are ORed together

Search Results

Name	Hometown
Bob Vance	Scranton, PA
Phyllis Lapin-Vance	Scranton, PA

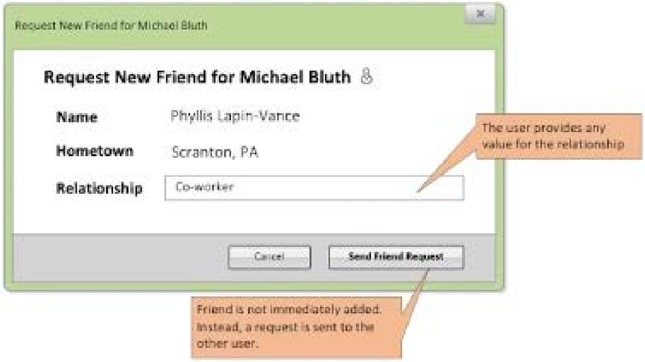
Existing friends should not be shown in the list

Clicking on a name shows the Add Friend dialog box

After clicking Search, any relevant results appear here (or on a new page)

Requirement: ...Social networks are all about making connections. GTOnline allows users to Search for Friends and Connect to them. There are several steps involved in making a connection with a new friend on GTOnline:

•The user searches for a friend based on several profile criteria including Name, Email Address, and Hometown.


Request New Friend for Michael Bluth

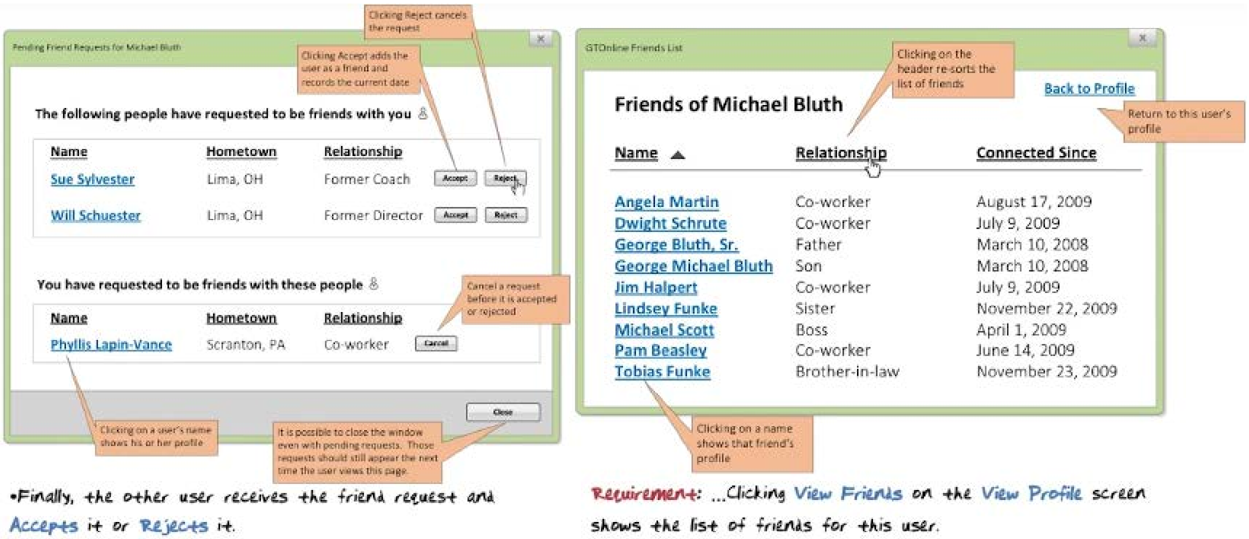
Request New Friend for Michael Bluth

Name	Phyllis Lapin-Vance
Hometown	Scranton, PA
Relationship	Co-worker

The user provides any value for the relationship

Friend is not immediately added. Instead, a request is sent to the other user.

•Then, the user submits a Friend Request to another user with whom they wish to connect.



Pending Friend Requests for Michael Bluth

The following people have requested to be friends with you

Name	Hometown	Relationship		
Sue Sylvester	Lima, OH	Former Coach	Accept	Reject
Will Schuester	Lima, OH	Former Director	Accept	Reject

You have requested to be friends with these people

Name	Hometown	Relationship	
Phyllis Lapin-Vance	Scranton, PA	Co-worker	Cancel

Close

Clicking Reject cancels the request

Clicking Accept adds the user as a friend and records the current date

Cancel a request before it is accepted or rejected

Clicking on a user's name shows his or her profile

It is possible to close the window even with pending requests. Those requests should still appear the next time the user views this page.

•Finally, the other user receives the friend request and Accepts it or Rejects it.

GTOnline Friends List

Friends of Michael Bluth

Back to Profile

Name ▲	Relationship	Connected Since
Angela Martin	Co-worker	August 17, 2009
Dwight Schrute	Co-worker	July 9, 2009
George Bluth, Sr.	Father	March 10, 2008
George Michael Bluth	Son	March 10, 2008
Jim Halpert	Co-worker	July 9, 2009
Lindsey Funke	Sister	November 22, 2009
Michael Scott	Boss	April 1, 2009
Pam Beasley	Co-worker	June 14, 2009
Tobias Funke	Brother-in-law	November 23, 2009

Clicking on the header re-sorts the list of friends

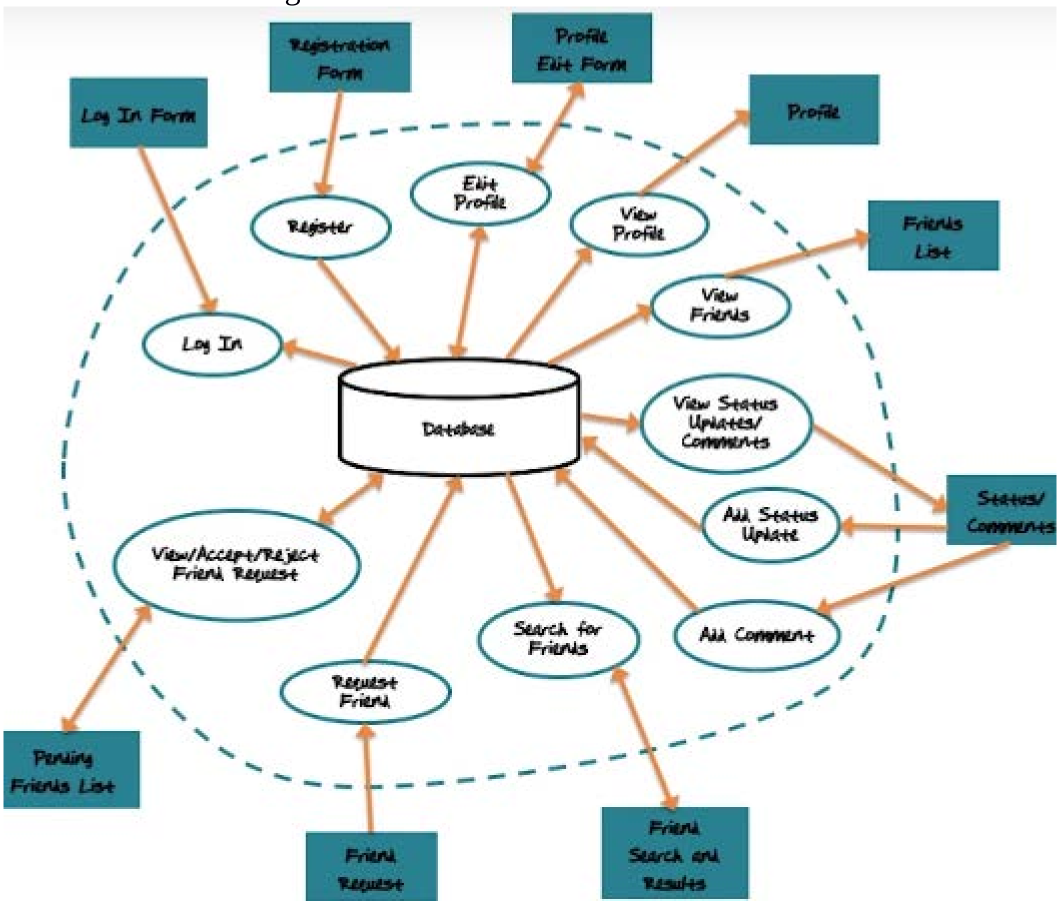
Return to this user's profile

Clicking on a name shows that friend's profile

Requirement: ...Clicking View Friends on the View Profile screen shows the list of friends for this user.

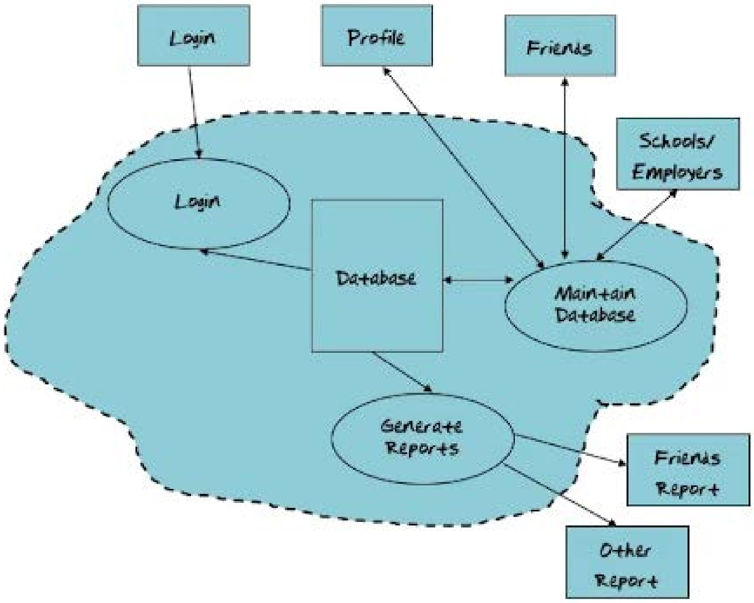
- This is an input AND output document
  - Input because of the buttons which can accept or reject
  - Output due to all other information displayed

Information Flow Diagram



- The output of the analysis phase is an information flow diagram
- Note the different tasks are represented with circles

- Note the different documents
- Note the arrows – they indicate the direction of the data / information flow
- The names of the documents and tasks very closely reflect that which is in the requirements document
- This is the first top level specification of the system – nothing more
- Note: DO NOT make this too generic – for example DO NOT CREATE THE IFD LIKE THIS:



- This does NOT model the future system we are building – this models programs and code, which is NOT what we want to do here

Next Phase: Specification

- We have looked at the customer requirements and designed an information flow diagram
- The next phase is that of specification; we are going to concentrate on specifying the database
    - To do this, we will look at the input / output documents to the system, and then design an extended entity relationship diagram
- We will then look at the tasks from the information flow diagram and will provide the specification of what the task will do relative to the input / output documents
- The output of the specification phase:
    - EER Diagram
    - Data Formats
    - Constraints
    - Task Decompression
- How does one go about designing the specification of the database?
    - Ask:
        - What goes into the database?
            - Everything in the database MUST come from somewhere

        - Everything on the input documents must go somewhere
    - What comes out of the database?
        - Everything in the database MUST be used for something
        - Everything on the output documents must come from somewhere
- With those questions answered we can develop the specifications for the database